Final Report for ETL Project

**Title: Correlation between Tourism and US-UK Exchange Rate.**

Extract Data:

**Data Source 1**: CSV file downloaded from <https://www.visitbritain.org/markets/usa>

The data source focuses on the US market detailing the number of US visitors to the UK per annual quarter with respect to duration stay, Air transport, nationality, purpose of visit, sex, age, spending and number of nights stayed.

The data ranges between 2002-2017.

**Data Source 2**: Web API for retrieval of US-UK exchange rate for 2000-2018. Used <https://fred.stlouisfed.org/> .

Transform Data:

In order to relate the US-UK exchange rate to the travel data, the exchange rate was averaged quarterly.

The travel data (.csv) file was read directly into a pandas dataframe and was then merged on the year and quarter values with the exchange rate values.

Load Data:

The pandas dataframe was then loaded into a Mongo DB. A non-relational database could provide more flexibility with searching queries.

The pandas dataframe was also loaded into a MySQL DB. The schema within MYSQL is probably more suitable than that in Mongo DB since the schema is set via the table columns and values.

Hypothetical Use of Database:

The database could be used to identify relationships between the US-UK exchange rate and the number of US ‘holiday’ visitors to the UK per annual quarter with respect to duration stay, Air transport, nationality, purpose of visit, sex, age, spending and number of nights stayed.